



DE LA RECHERCHE À L'INDUSTRIE

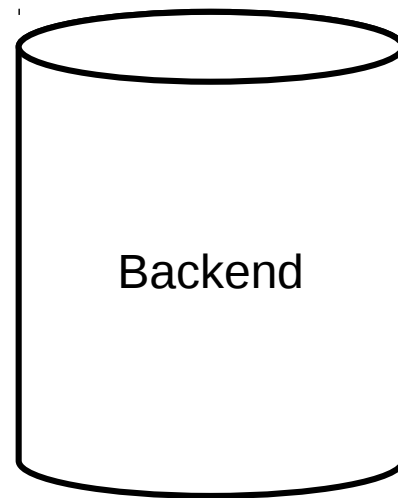
# RobinHood-v4 progress report

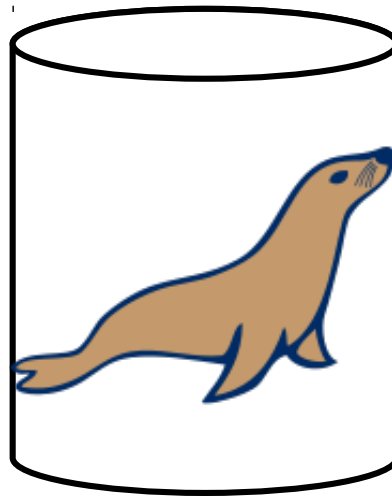
Quentin Bouget

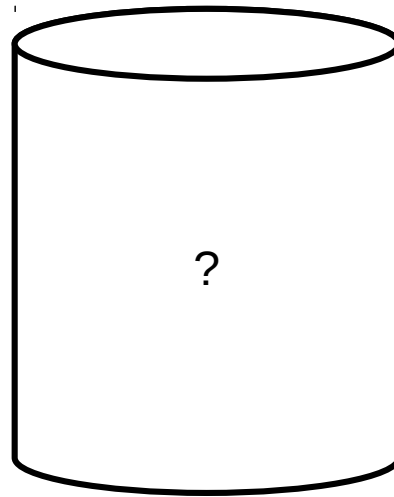
# What is RobinHood?

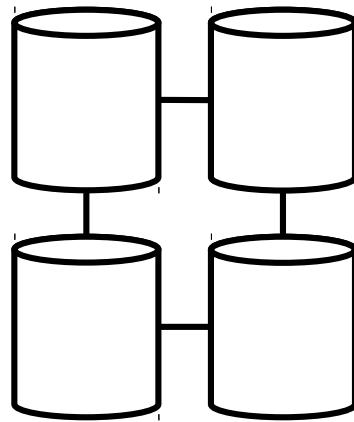
RobinHood mirrors a filesystem's metadata,  
and makes it queryable

- Mirror
  - Read Replica
  - Snapshot
  - Disaster Recovery
- Query
  - Expressiveness
  - Performance
  - Efficiency

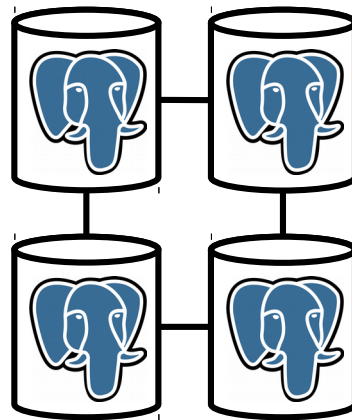


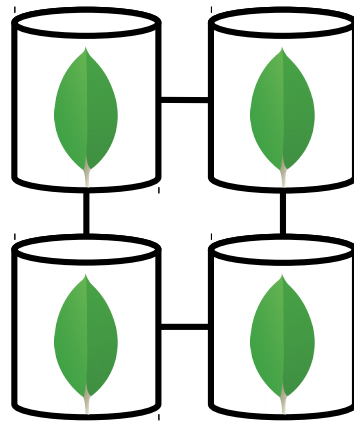


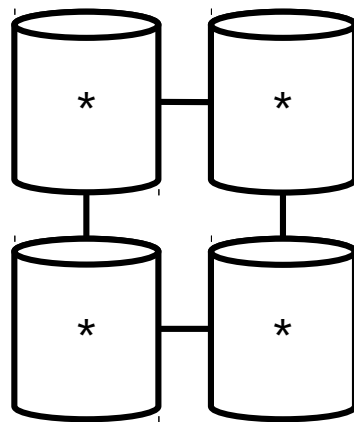


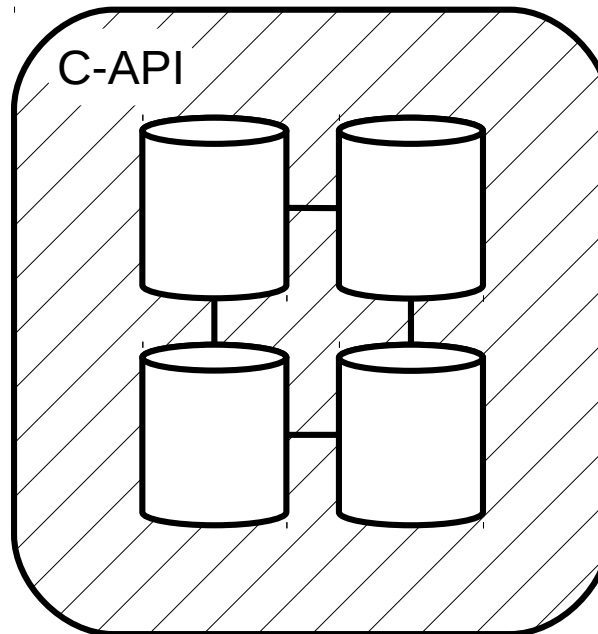


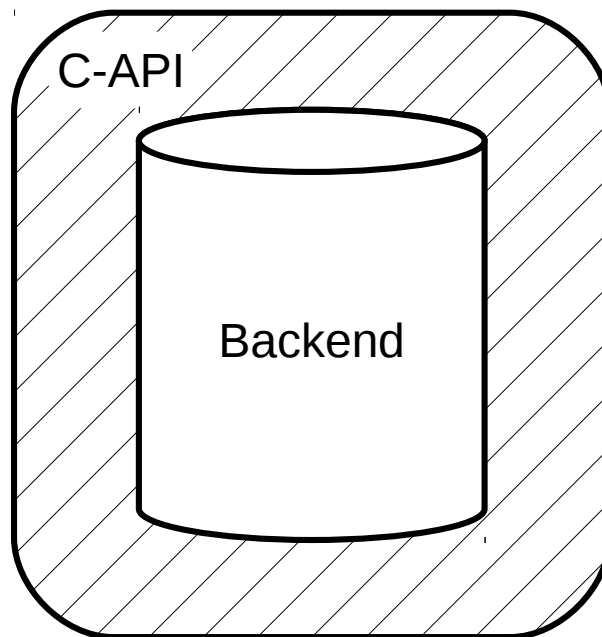






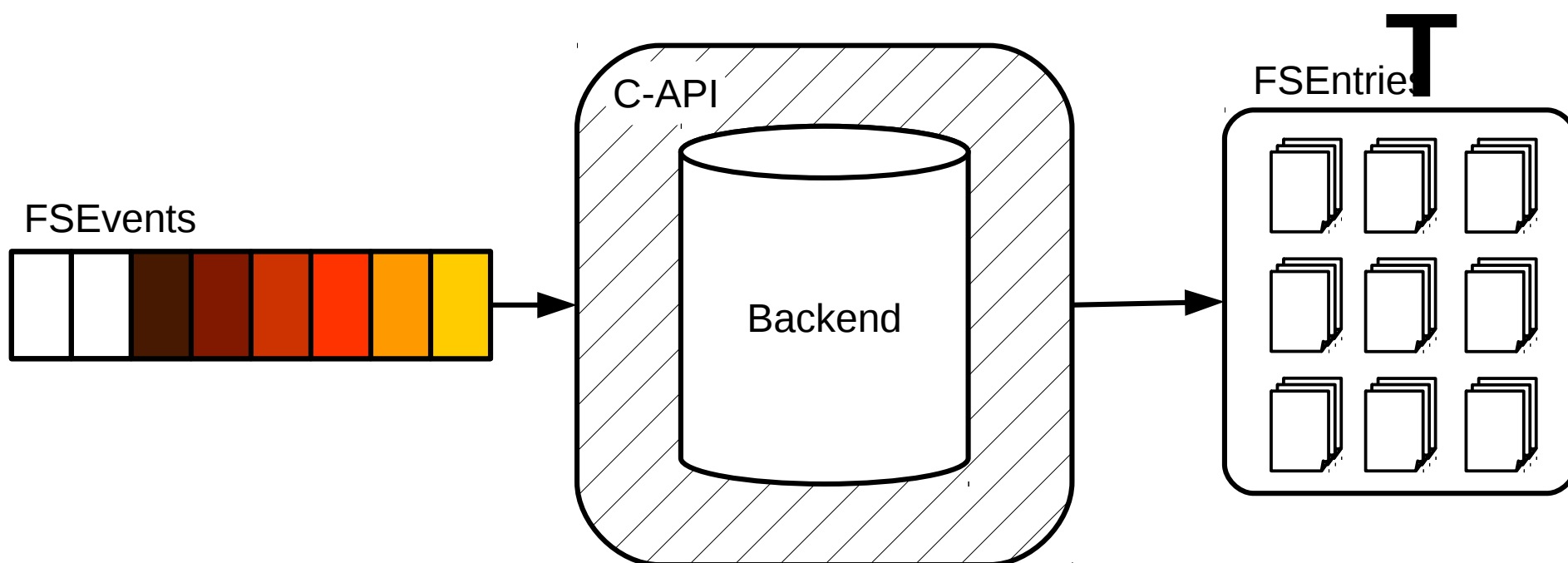






IN

OU



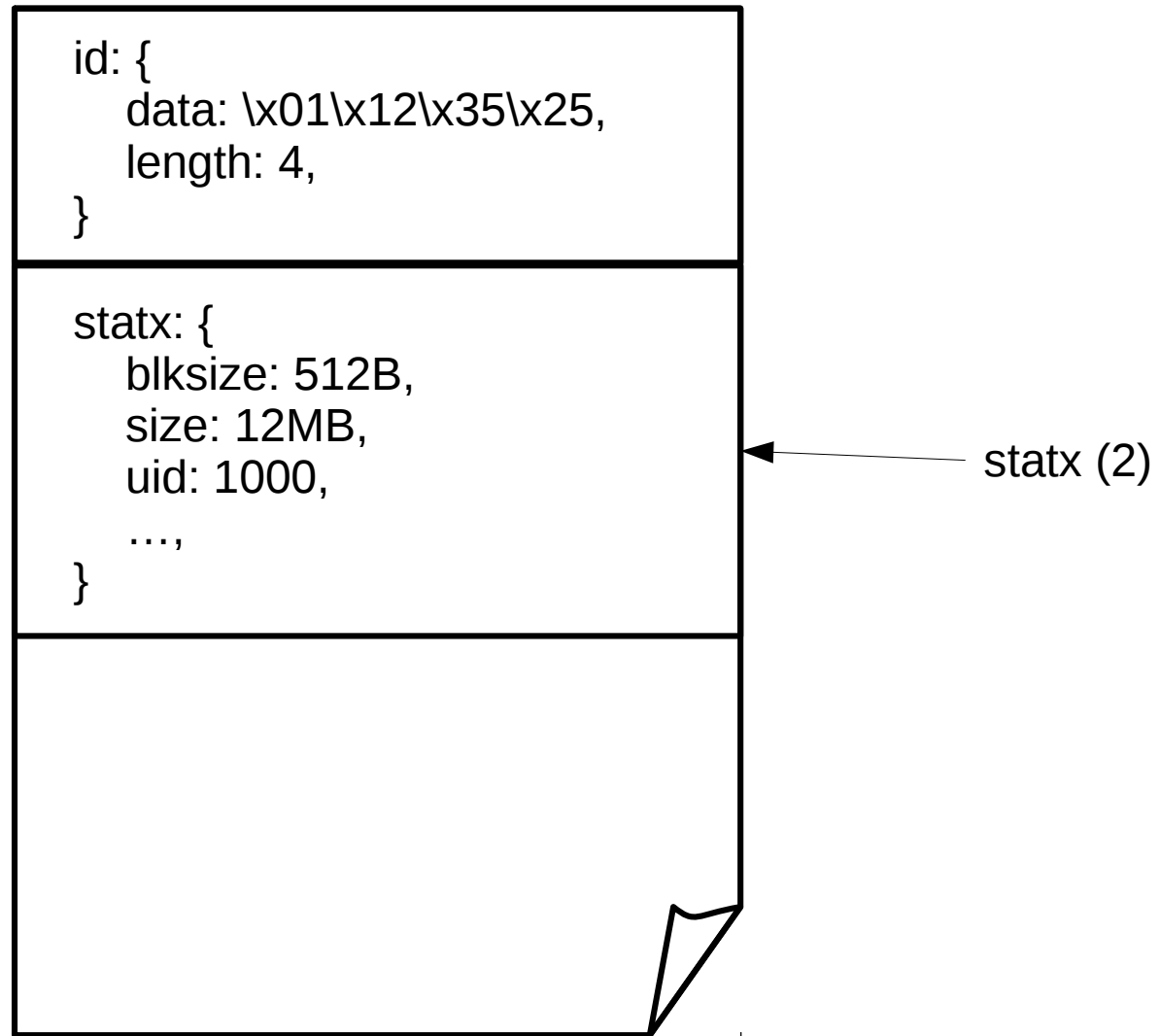


# FSEntry

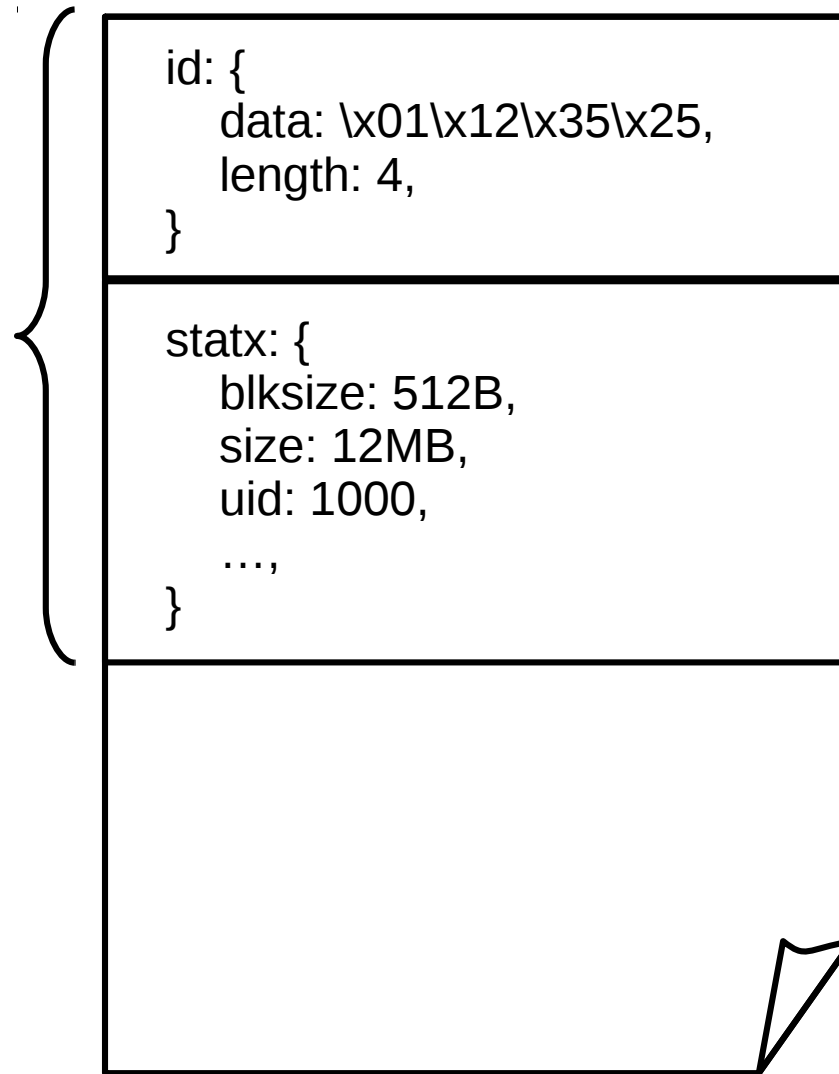


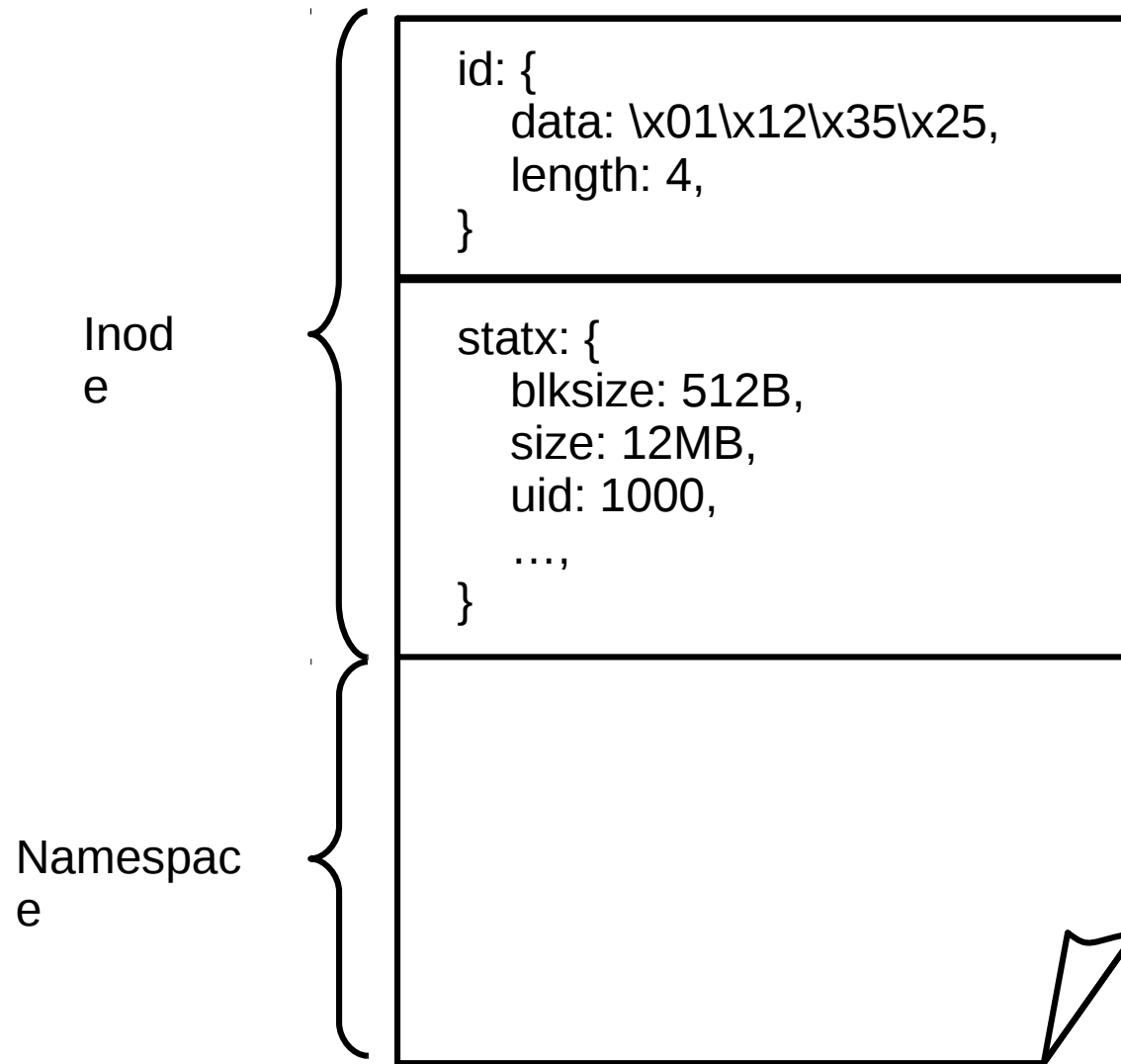


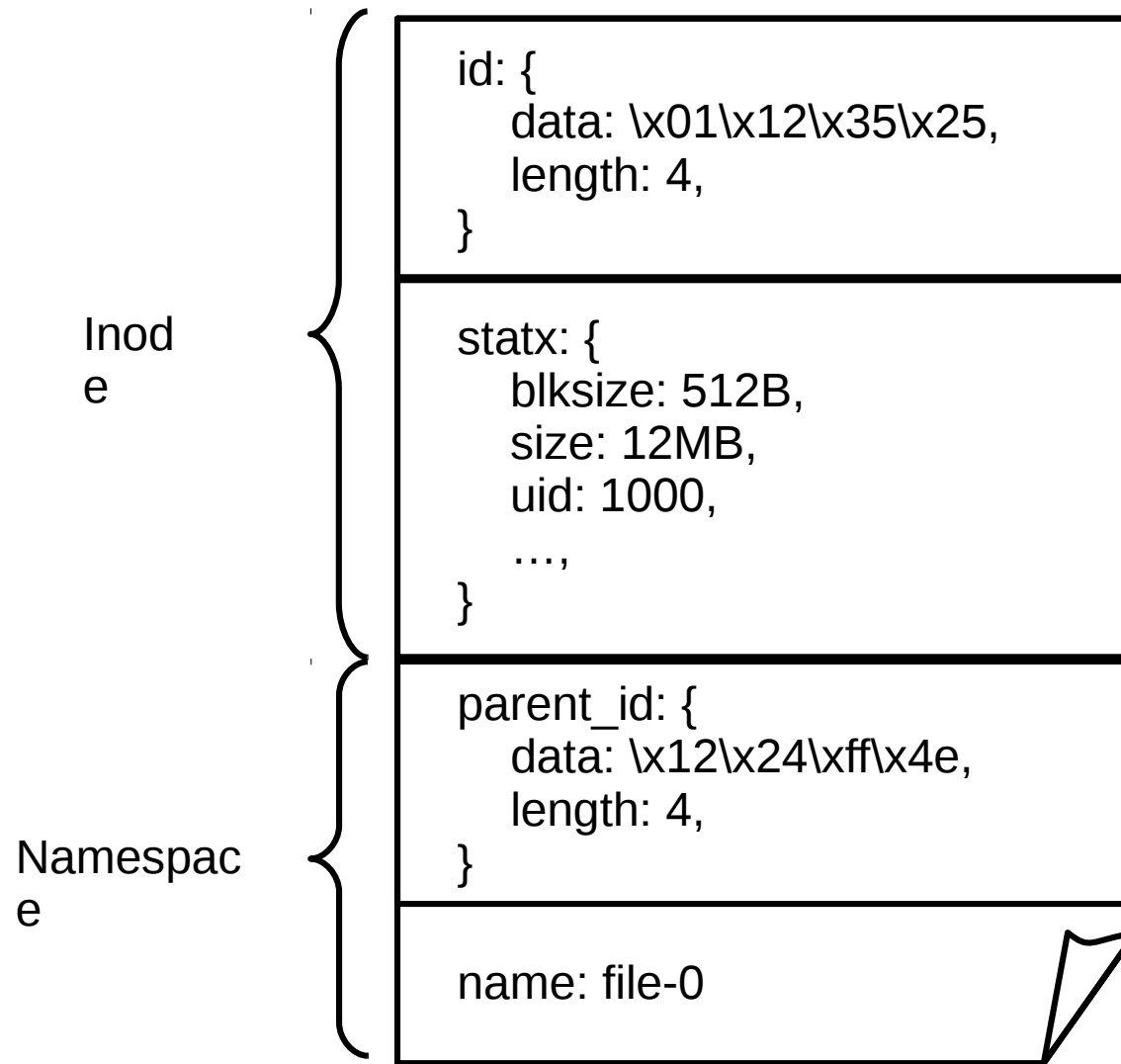




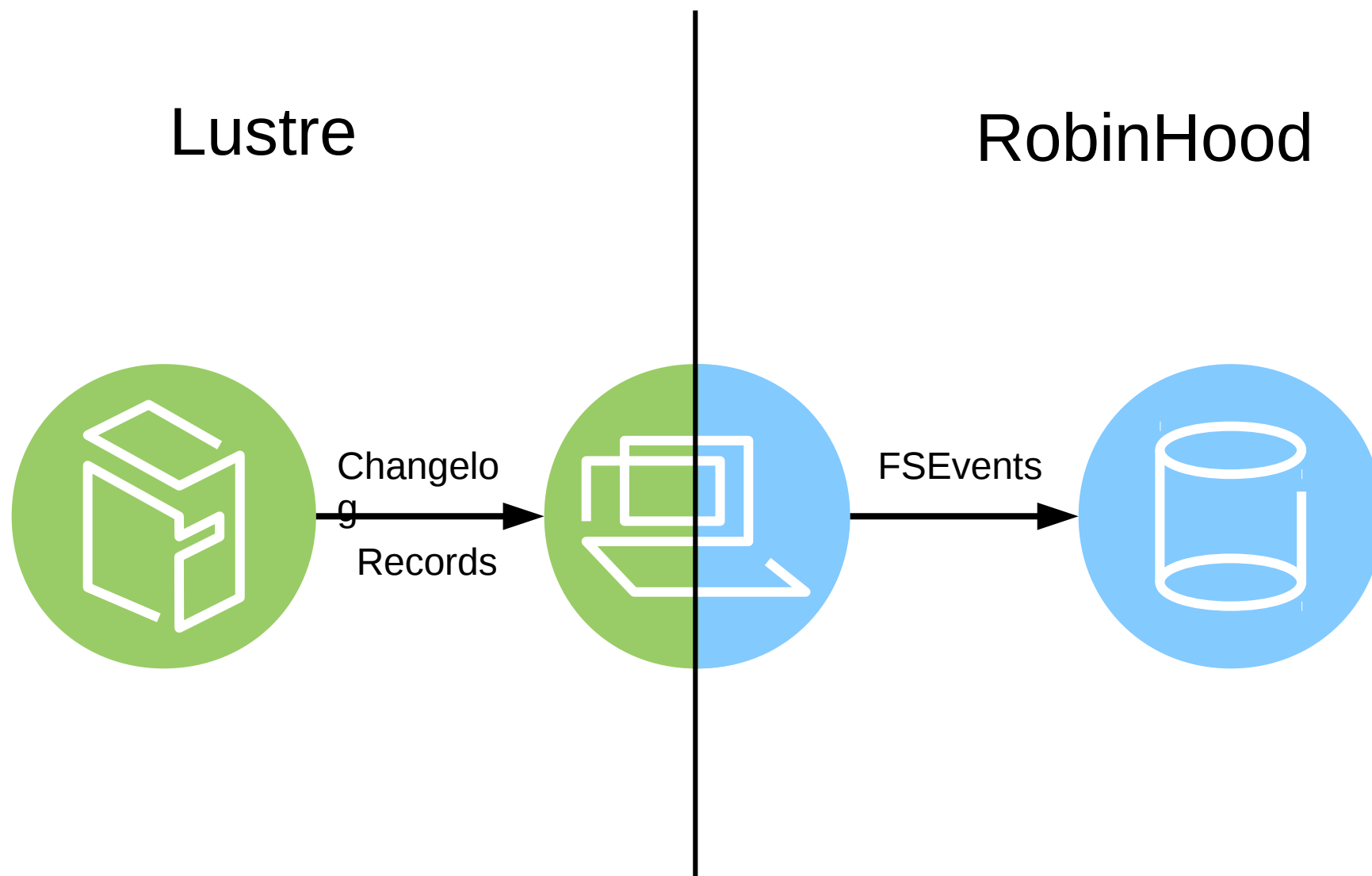
Inode







# FSEvent



```
$> touch file-0
```

```
$> rm file-0
```

```
$> lfs changelog lustre-MDT0000
```

```
CREAT t=[FID#1] p=[FID#0] file-0
```

```
UNLNK t=[FID#1] p=[FID#0] file-0
```



# CREA T

[FID#1]
_____ _____ _____
[FID#0]
file-0

# UNLN K

[FID#1]
_____ _____ _____
[FID#0]
file-0

CREAT
[FID#1]
_____ _____ _____
[FID#0]
file-0

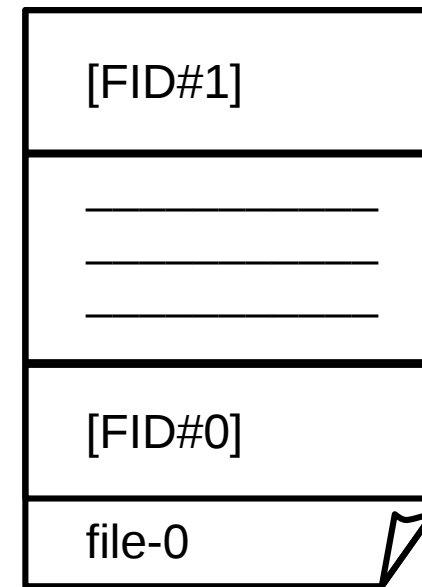
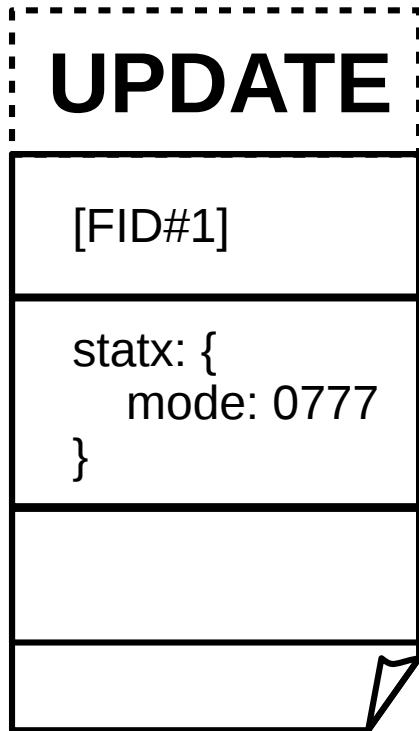
UNLNK
[FID#1]
_____ _____ _____
_____ _____ _____
_____

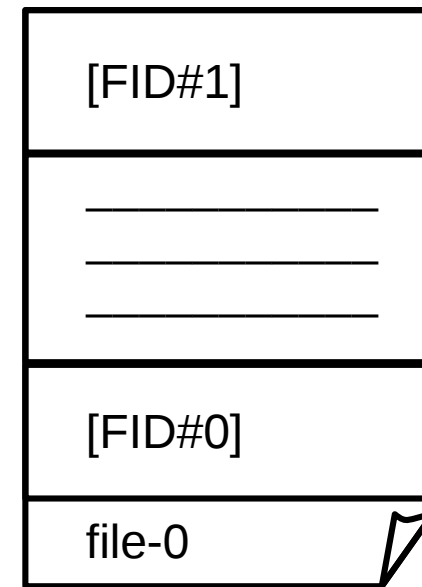
INSERT
[FID#1]
_____ _____ _____
[FID#0]
file-0

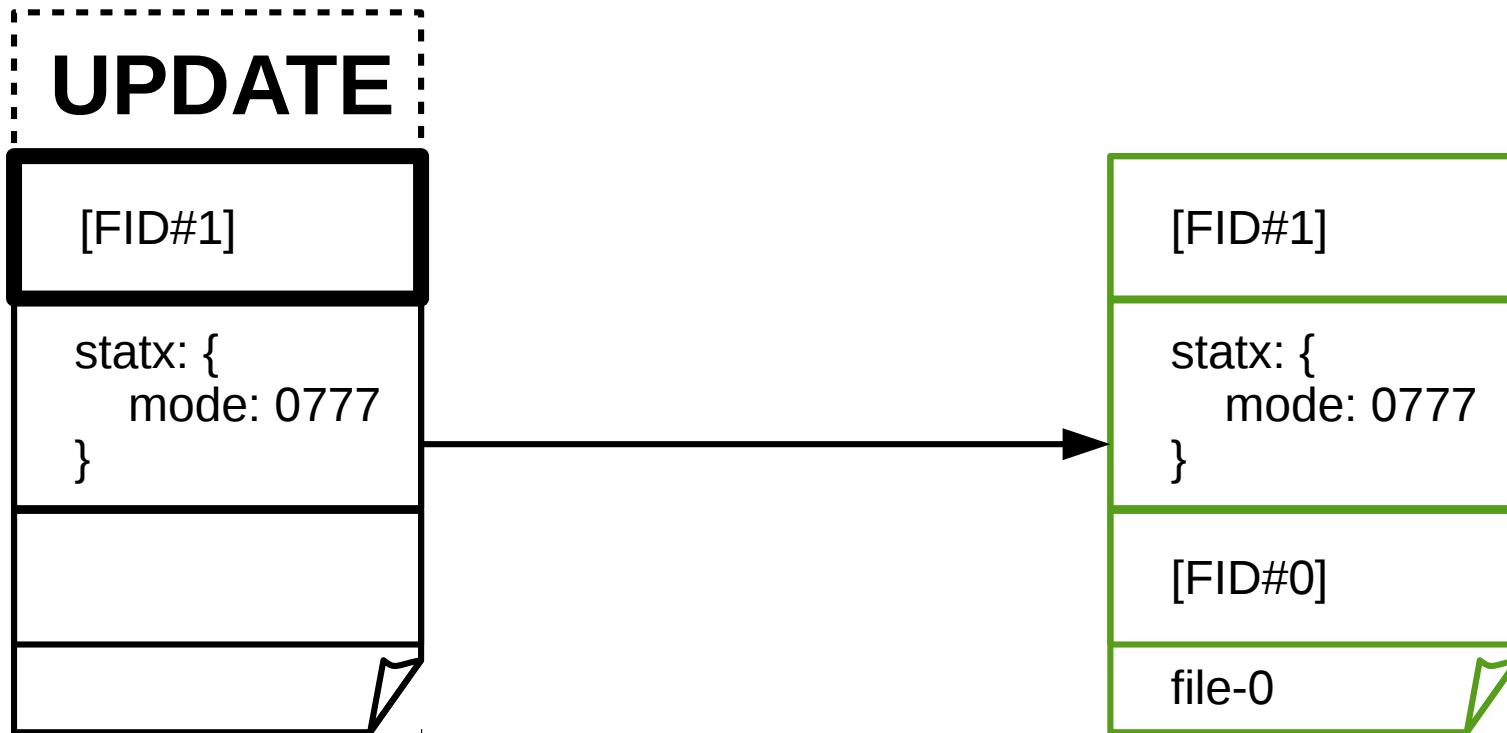
DELETE
[FID#1]
_____ _____ _____
_____
_____

```
$> chmod a+rwx file-0
```

```
$> lfs changelog lustre-MDT0000  
SATTR t=[FID#1]
```



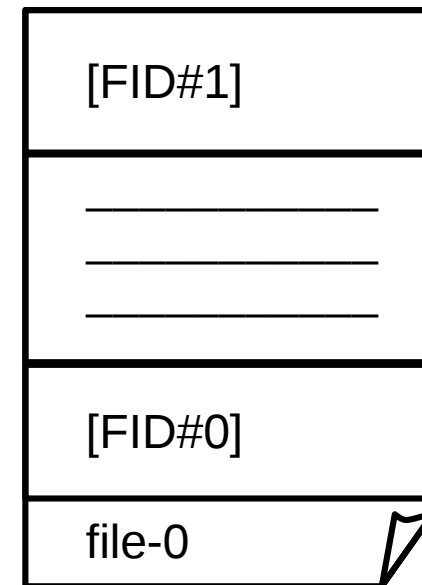




```
$> mv file-0 file-1
```

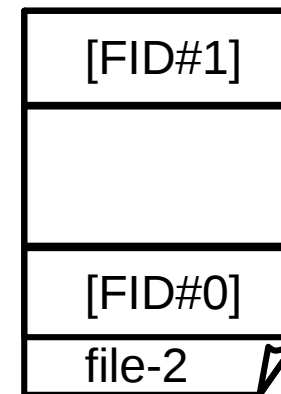
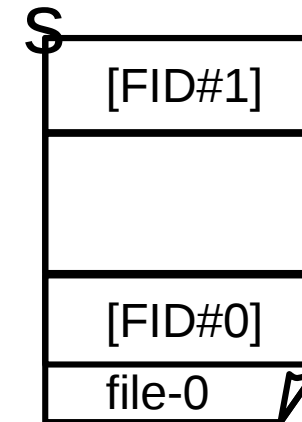
```
$> lfs changelog lustre-MDT00000  
RENME t=0 p=[FID#0] file-1 s=[FID#1]  
sp=[FID#0] file-0
```

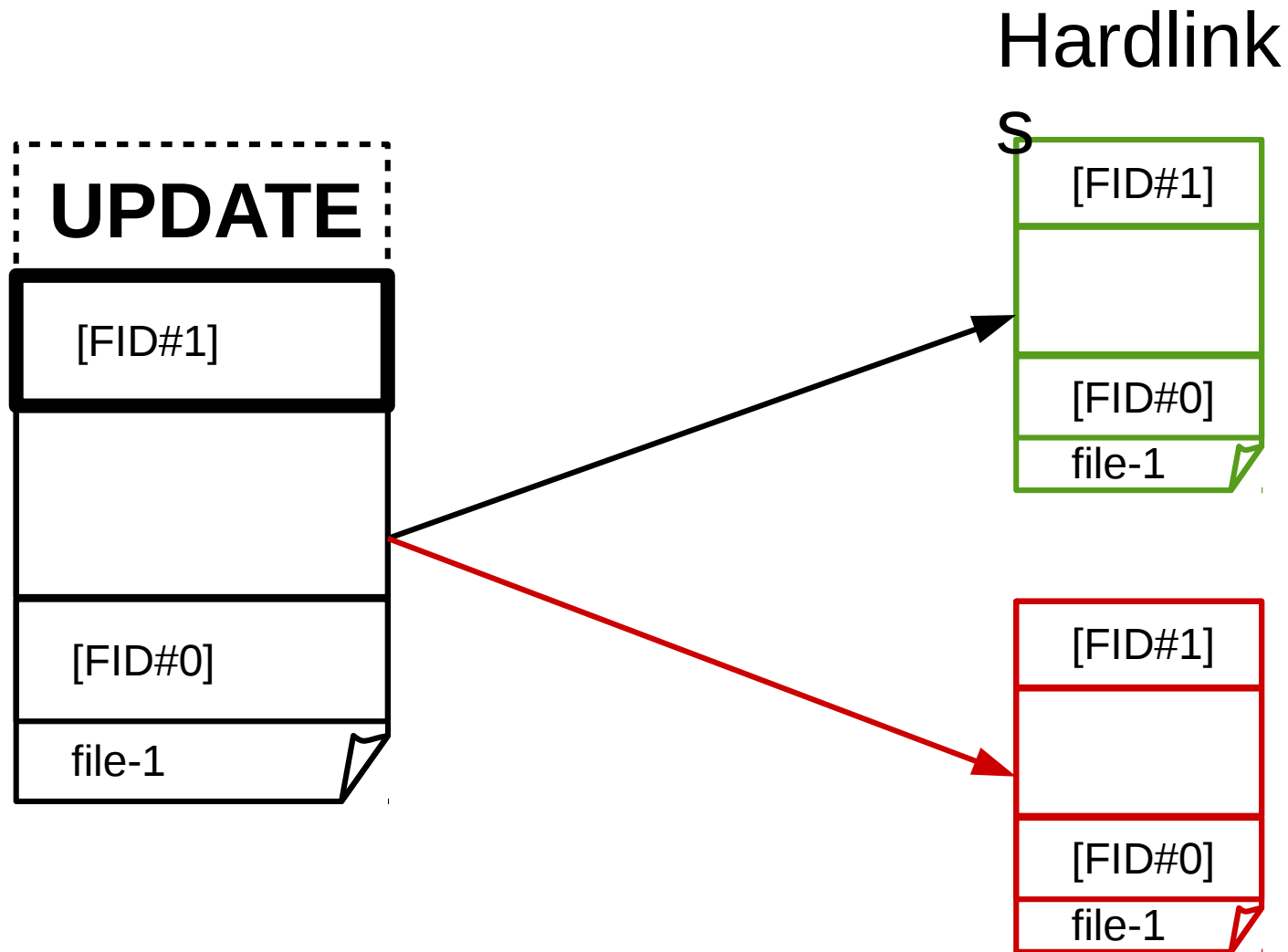


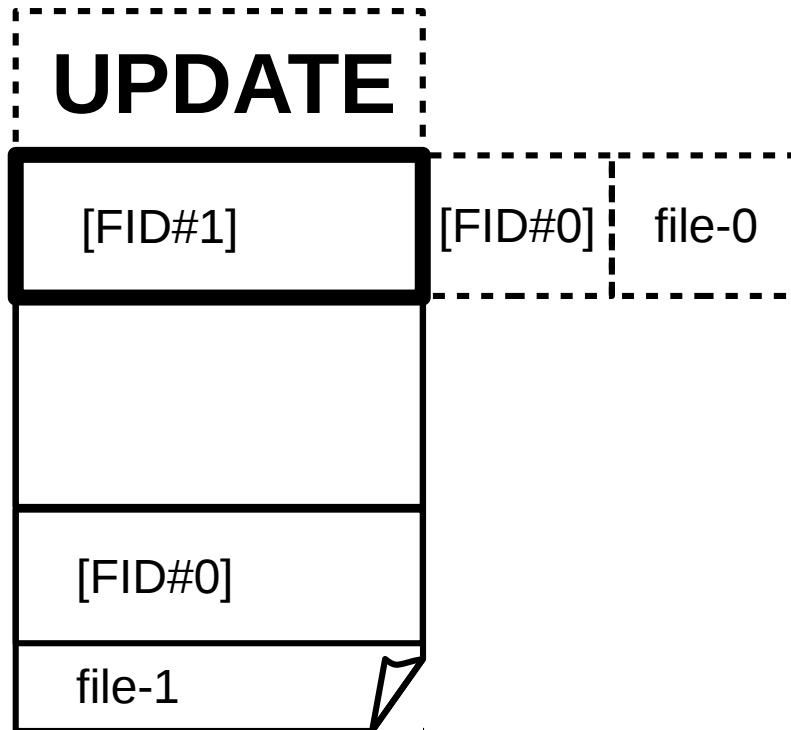




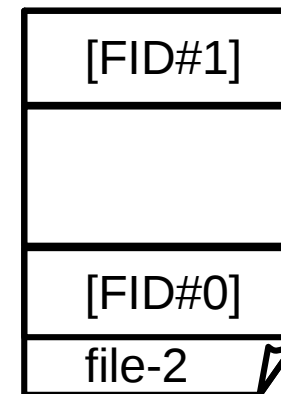
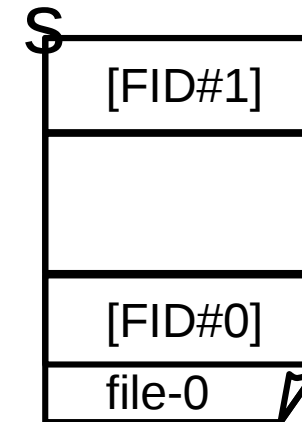
## Hardlink

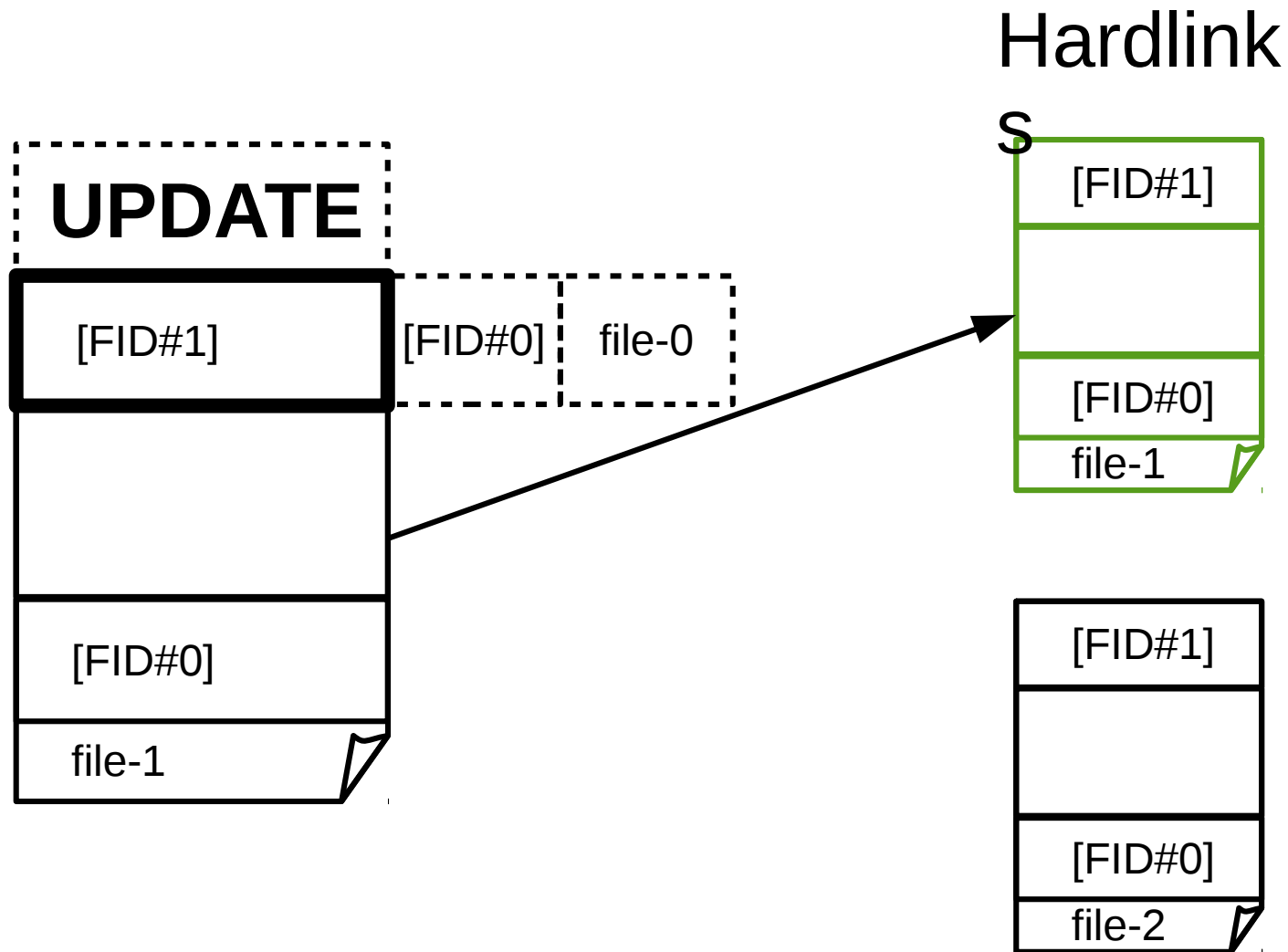


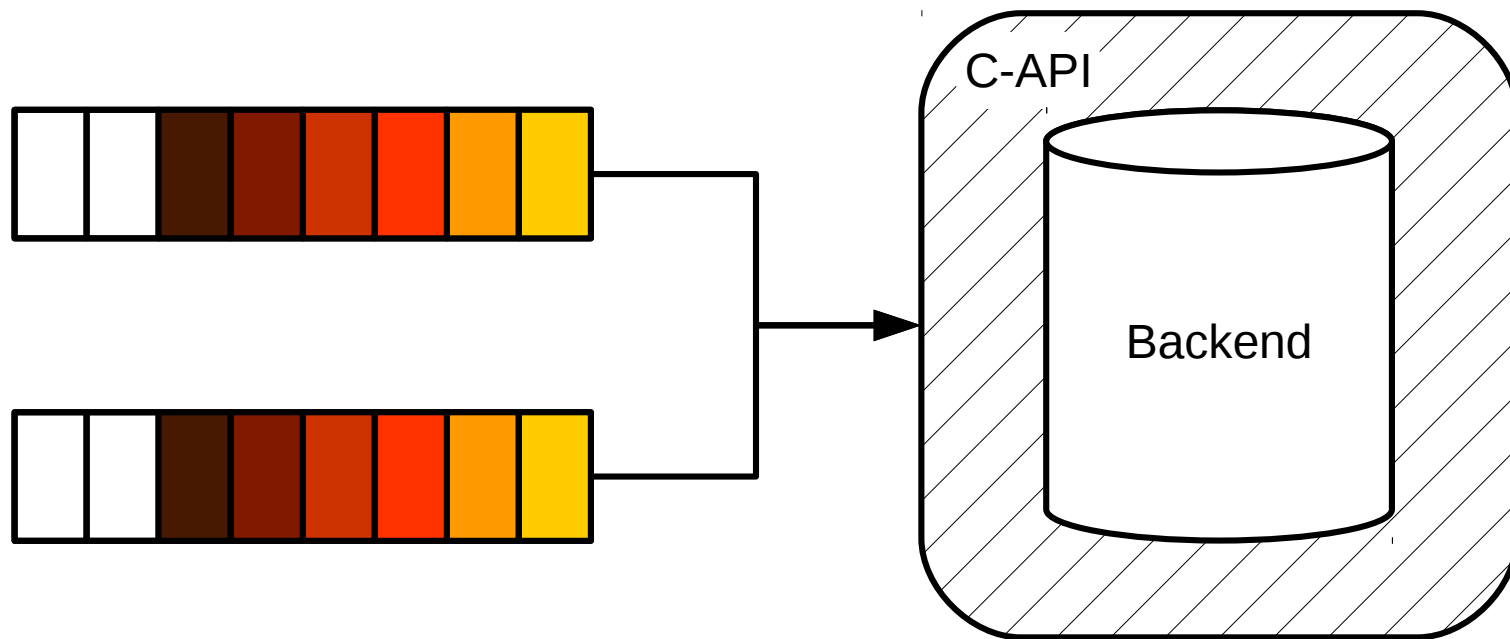




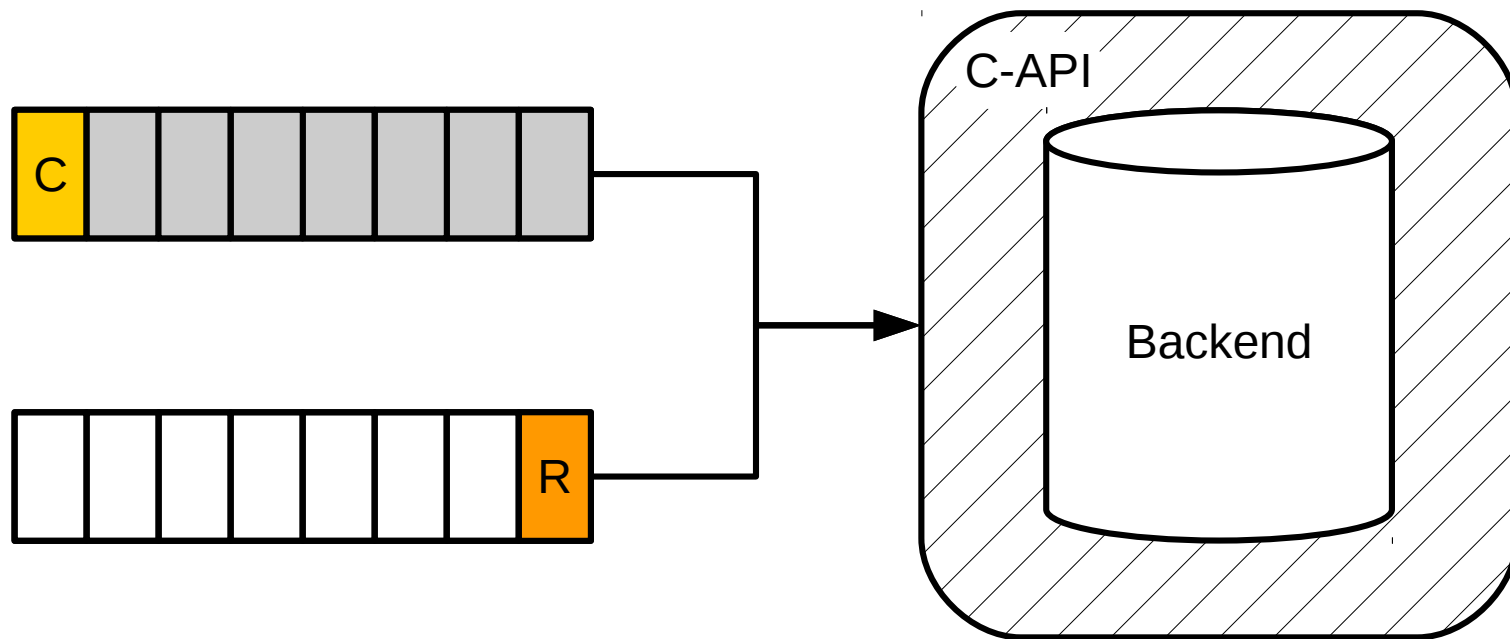
## Hardlink



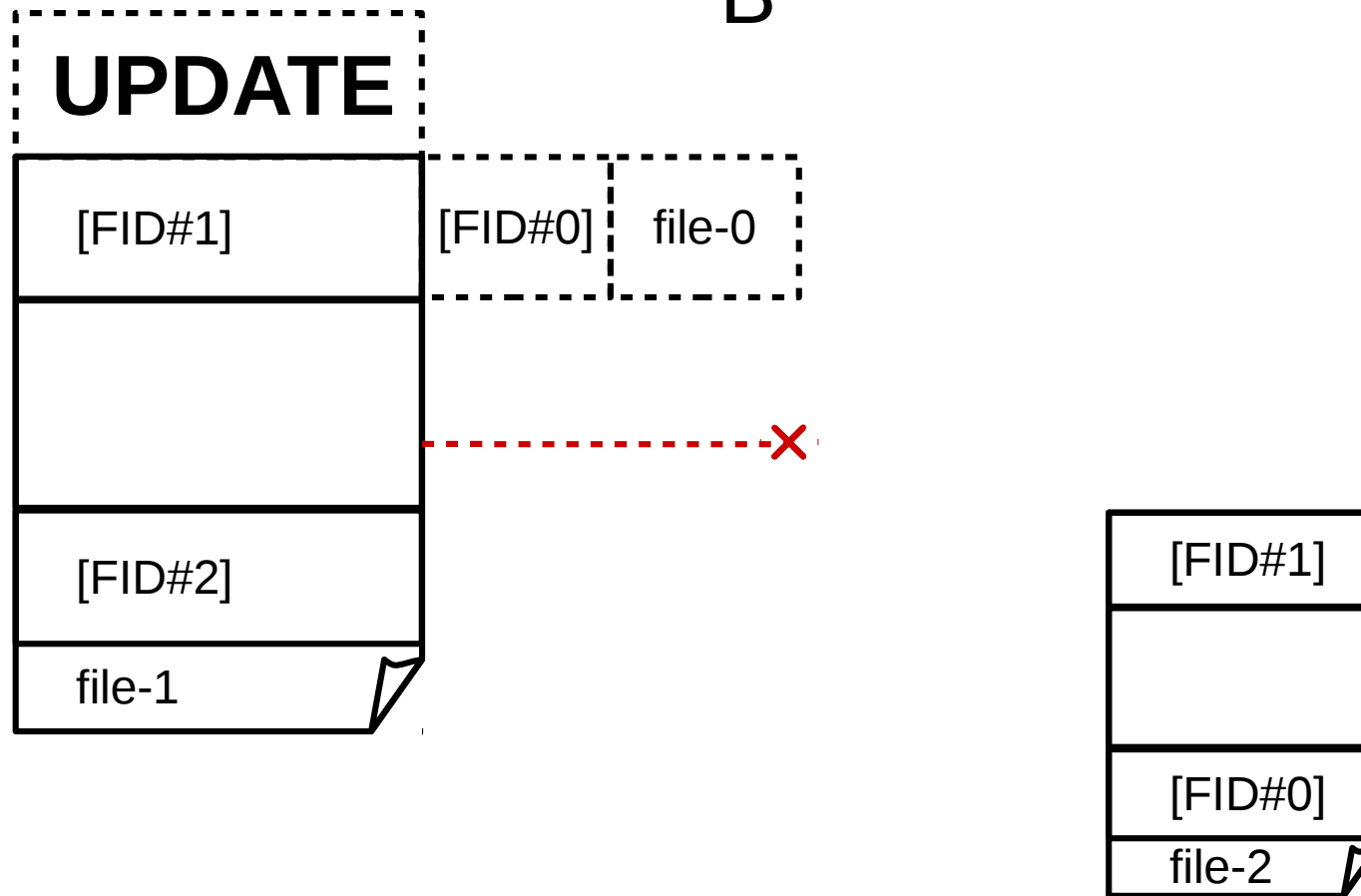




# Hardlinks + DNEv2 phase B

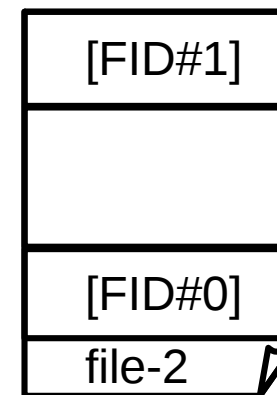
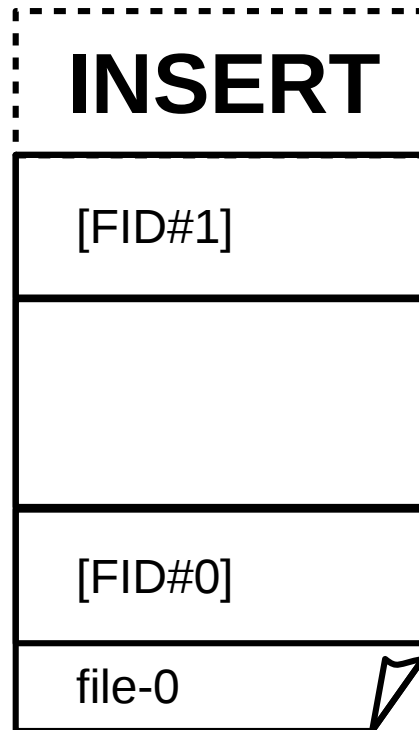


# Hardlinks + DNEv2 phase B

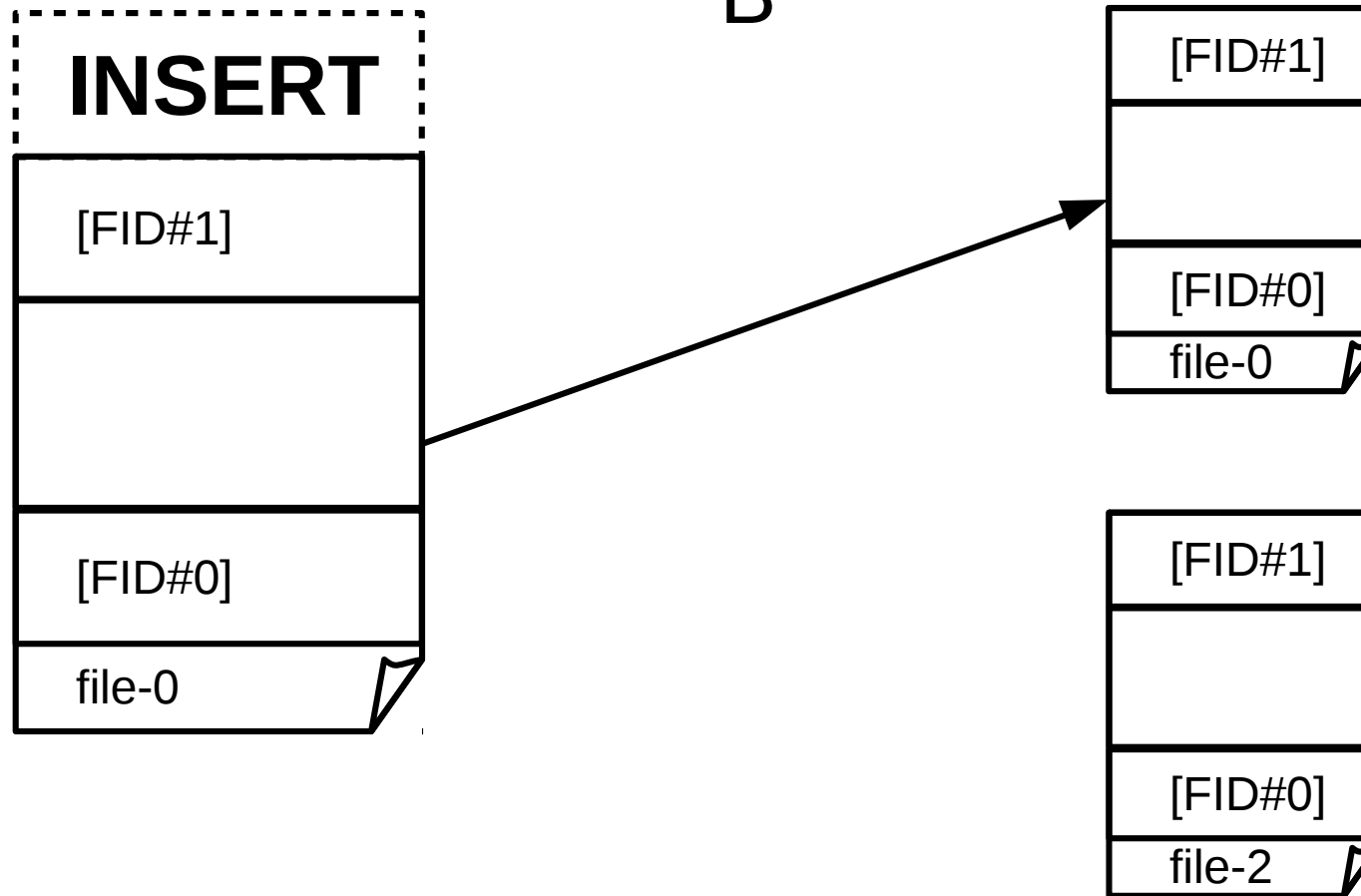




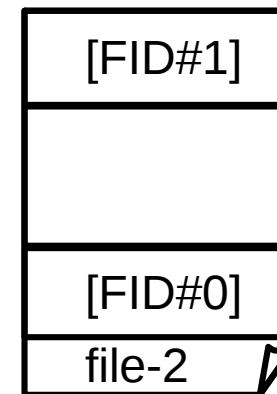
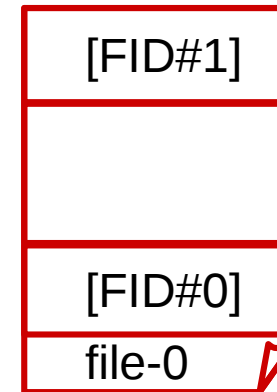
## Hardlinks + DNEv2 phase B



## Hardlinks + DNEv2 phase B

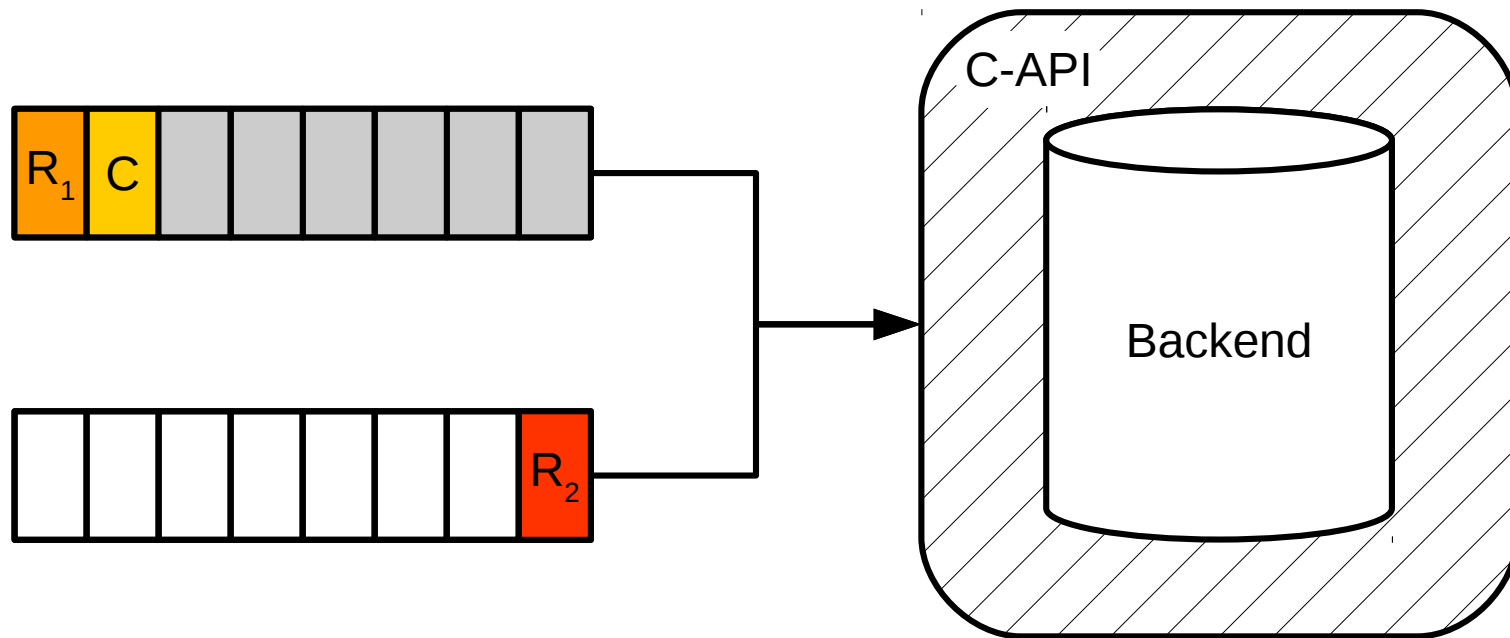


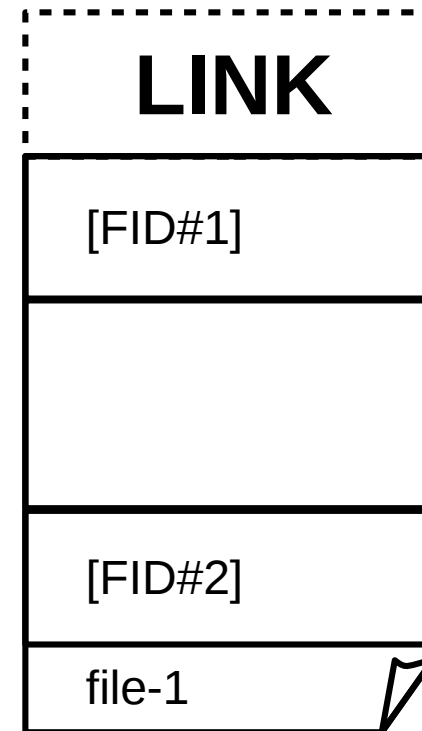
# Hardlinks + DNEv2 phase B

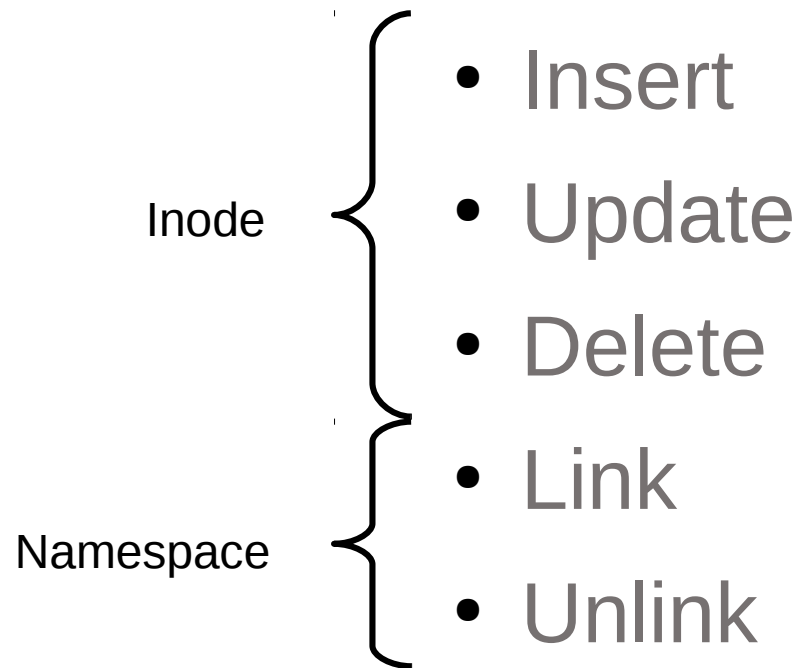


LU-12754

# Hardlinks + DNEv2 phase B







# Applications



rbh-sync

# Synchronize two RobinHood backends

# rbh-sync SOURCE DEST

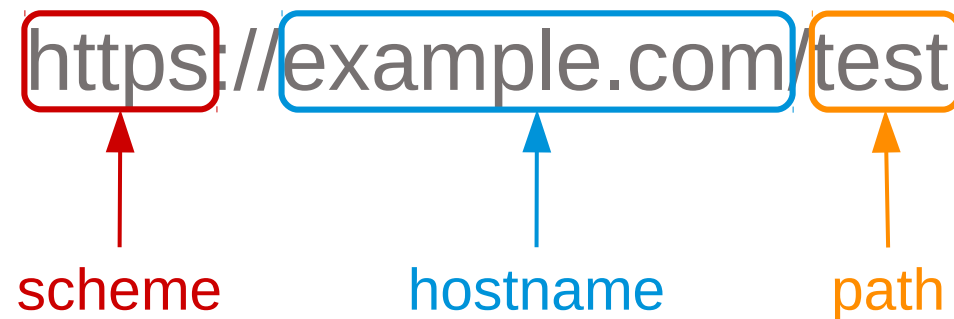
- Uniform Resource Identifier
- RFC 3986:

scheme:[//[userinfo@]host[:port]]path

- Uniform Resource Identifier
- RFC 3986:

scheme:[//[userinfo@]host[:port]]path  
[?query][#fragment]

- Uniform Resource Identifier
- RFC 3986:

The diagram shows the URI 'https://example.com/test' with three components highlighted by colored boxes: 'https' in a red box, 'example.com' in a blue box, and 'test' in an orange box. Below each box is a colored arrow pointing to it, with a label in the same color: a red arrow labeled 'scheme' points to 'https', a blue arrow labeled 'hostname' points to 'example.com', and an orange arrow labeled 'path' points to 'test'.

https://example.com/test

scheme hostname path

- Uniform Resource Identifier
- RFC 3986:

rbh:mongo:scratch

The diagram shows the URI 'rbh:mongo:scratch' with two labels and arrows pointing to its parts. A red label 'scheme' has a red arrow pointing to the 'rbh' part, which is enclosed in a red rounded rectangle. An orange label 'path' has an orange arrow pointing to the 'mongo:scratch' part, which is enclosed in an orange rounded rectangle.

- Uniform Resource Identifier
- RFC 3986:

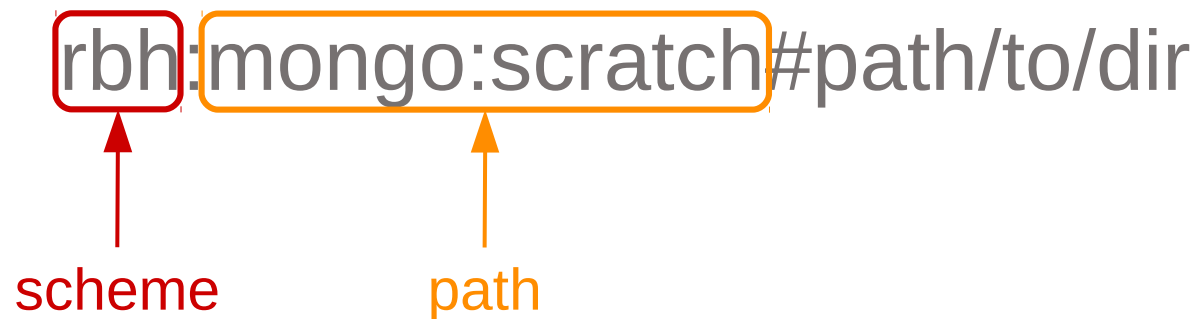
The diagram shows the URI 'rbh:mongo:scratch' with three components highlighted in colored boxes: 'rbh' in a red box, 'mongo' in a yellow box, and 'scratch' in a purple box. Arrows point from labels below to each component: a red arrow from 'scheme' to 'rbh', a yellow arrow from 'backend-type' to 'mongo', and a purple arrow from 'fsname' to 'scratch'.

rbh:mongo:scratch

scheme backend-type fsname



- Uniform Resource Identifier
- RFC 3986:

The diagram shows the URI 'rbh:mongo:scratch#path/to/dir'. The 'rbh' part is enclosed in a red rounded rectangle, with a red arrow pointing down to the word 'scheme' in red text. The 'mongo:scratch' part is enclosed in an orange rounded rectangle, with an orange arrow pointing down to the word 'path' in orange text. The '#' and 'path/to/dir' part is not enclosed in a box.

rbh:mongo:scratch#path/to/dir

scheme path

- Uniform Resource Identifier
- RFC 3986:

`rbh:mongo:scratch#path/to/dir`

The diagram illustrates the components of the URI `rbh:mongo:scratch#path/to/dir`. The `rbh` part is enclosed in a red box, with a red arrow pointing to the label `scheme` below it. The `mongo:scratch` part is enclosed in an orange box, with an orange arrow pointing to the label `path` below it. The `#path/to/dir` part is enclosed in a blue box, with a blue arrow pointing to the label `fragment` below it.

- Uniform Resource Identifier
- RFC 3986:

`rbh:mongo:scratch#path/to/dir`

The diagram illustrates the components of the URI `rbh:mongo:scratch#path/to/dir`. The `rbh` part is enclosed in a red box with a red arrow pointing to the label `scheme` below it. The `mongo:scratch` part is enclosed in an orange box with an orange arrow pointing to the label `path` below it. The `#path/to/dir` part is enclosed in a blue box with a blue arrow pointing to the label `path` below it.

- Uniform Resource Identifier
- RFC 3986:

`rbh:mongo:scratch#[0x2000000007:0x1:0x0]`



The diagram shows the URI `rbh:mongo:scratch#[0x2000000007:0x1:0x0]`. The `rbh` part is enclosed in a red box, and the `mongo:scratch` part is enclosed in an orange box. A red arrow points from the word `scheme` below to the `rbh` box. An orange arrow points from the word `path` below to the `mongo:scratch` box.

scheme

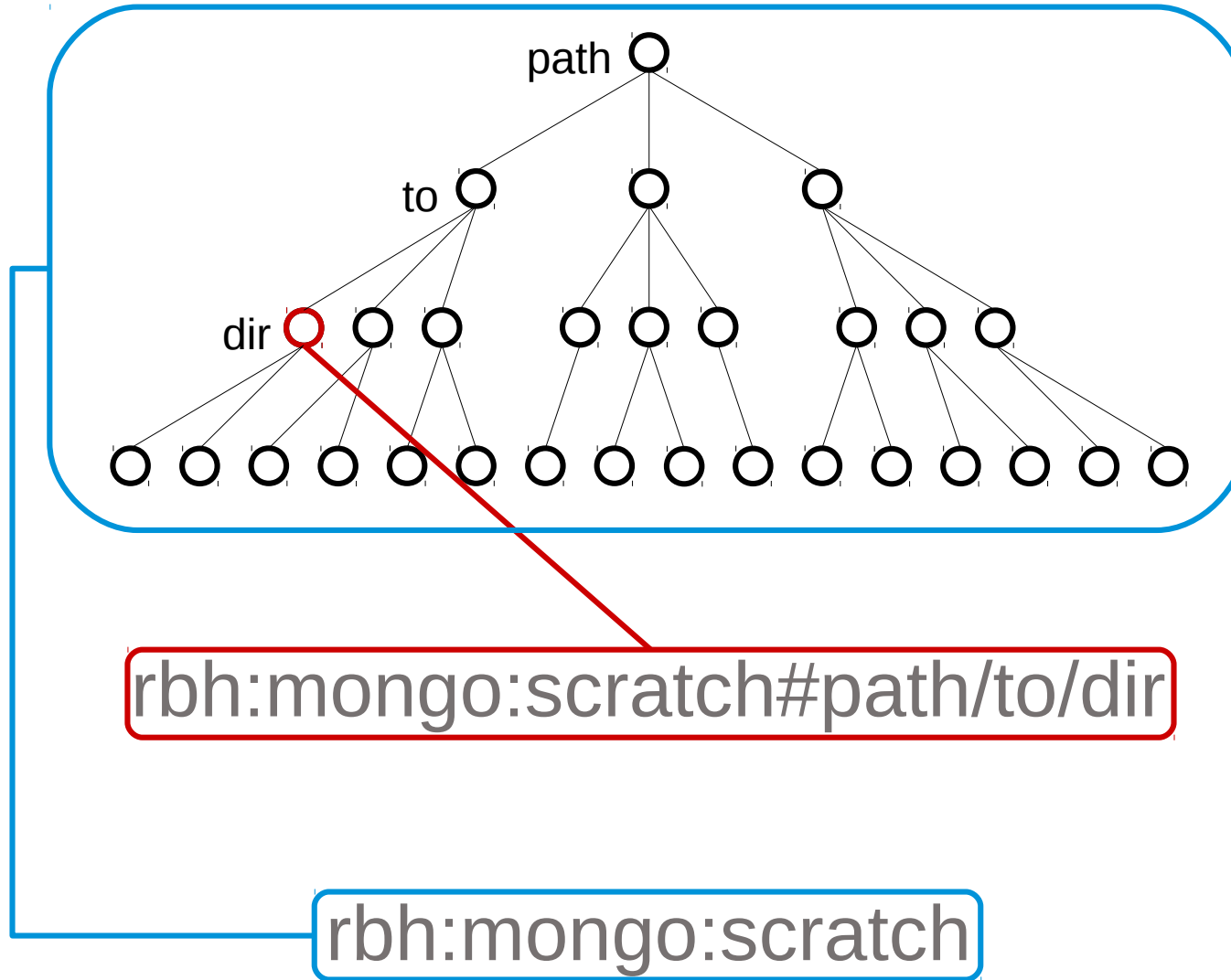
path

- Uniform Resource Identifier
- RFC 3986:

`rbh:mongo:scratch#[0x2000000007:0x1:0x0]`

The diagram shows the URI `rbh:mongo:scratch#[0x2000000007:0x1:0x0]` with three components highlighted by colored boxes and labeled with arrows below them:

- rbh**: Highlighted with a red box, labeled **scheme** with a red arrow.
- mongo:scratch**: Highlighted with an orange box, labeled **path** with an orange arrow.
- [0x2000000007:0x1:0x0]**: Highlighted with a blue box, labeled **FID** with a blue arrow.



rbh-sync posix:/mnt/scratch mongo:scratch

```
cd /mnt/scratch
for dir in *; do
    rbh-sync posix:/mnt/scratch#"$dir" mongo:scratch &
done
wait
```



- 1 node:
    - 16 cores
  - Lustre on RAM:
    - 1 MDT
  - 1M files + directories
- <GRAPH>
    - rbh-v3
    - rbh-v4
    - find -ls

rbh-find

A clone of (gnu-)find that queries RobinHood

rbh-find [URI ...] [[predicates] [actions] ...]

```
rbh-find mongo:scratch -name 'a*' -print0
```

- 1 node:
    - 16 cores
  - Lustre on RAM:
    - 1 MDT
  - 1M files + directories
- <GRAPH>
    - find
    - lfs find
    - rbh-find-v3
    - rbh-find-v4
  - + rbh-v3 → ∅ path
  - + lfs find → uid-only

rbh-fsevent

# Apply filesystem events on a RobinHood backend

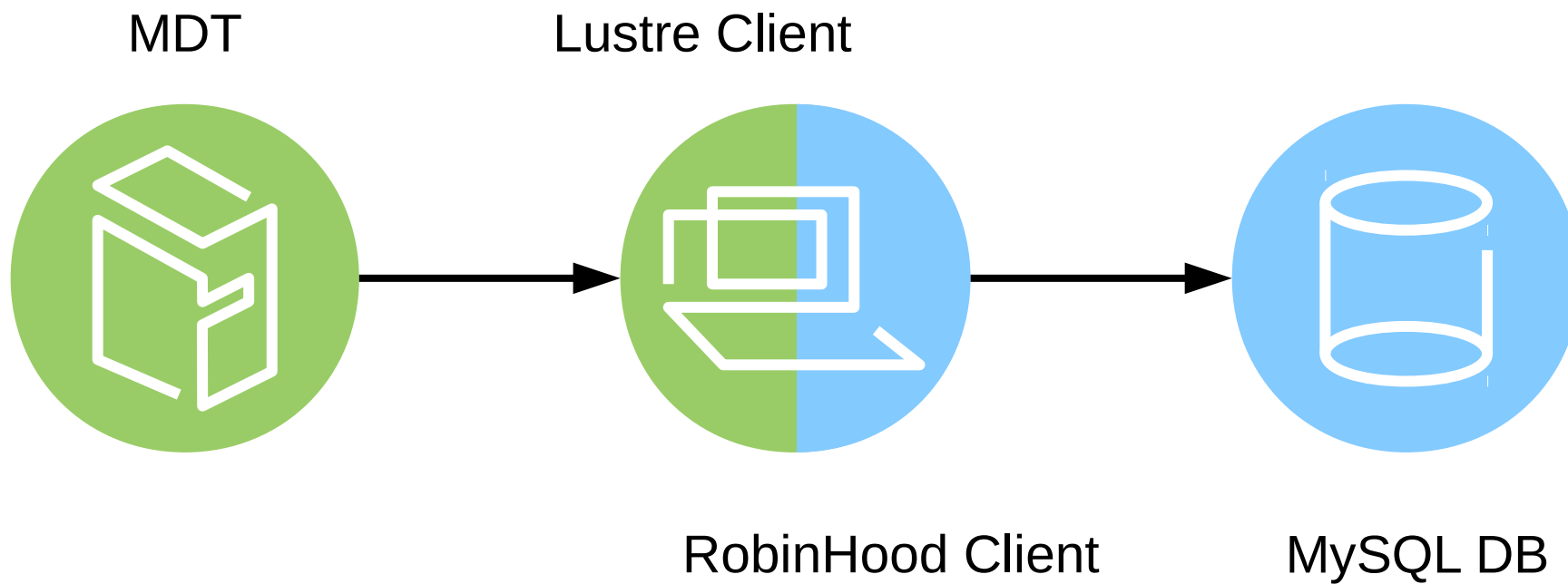


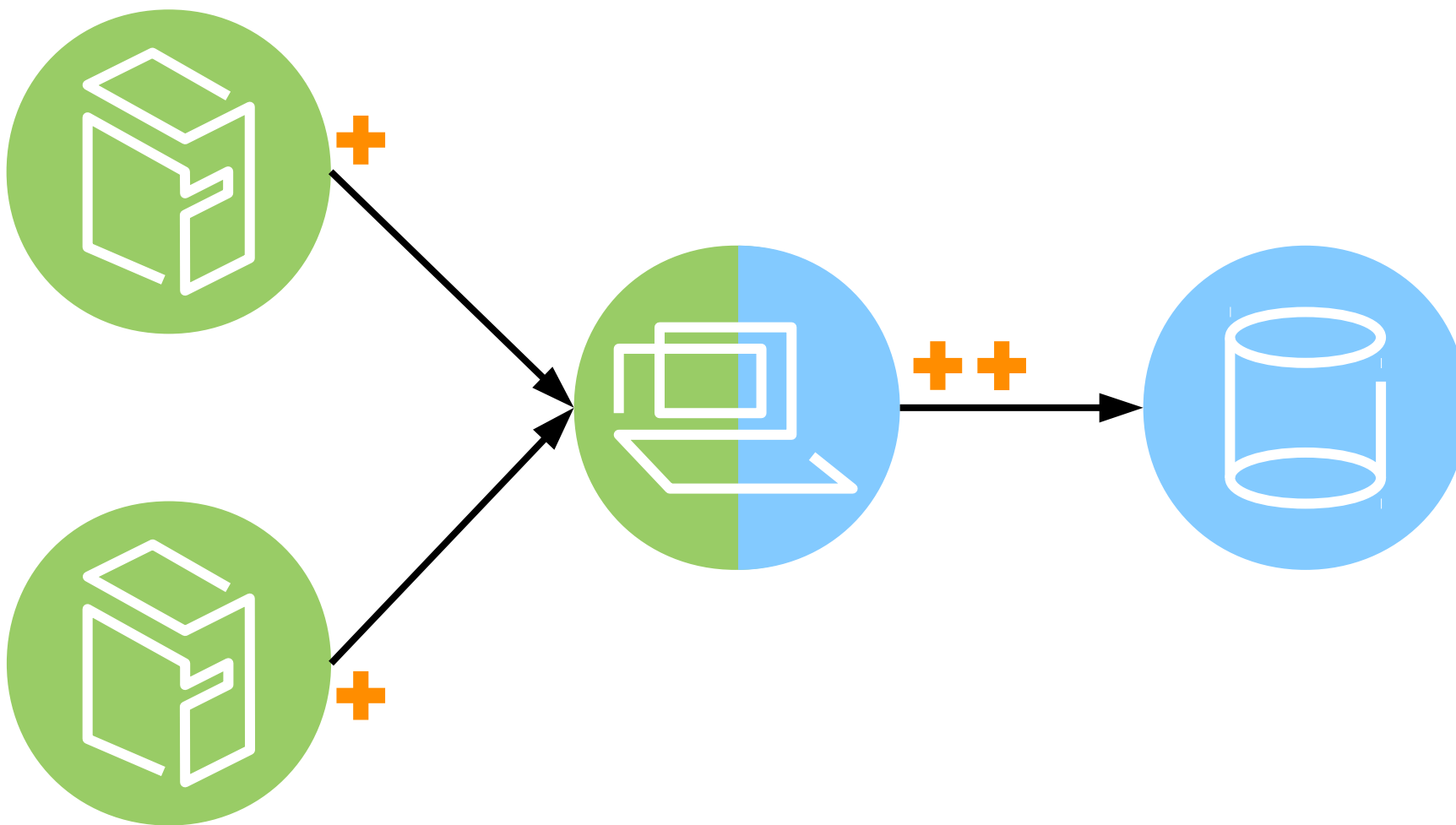
TODO

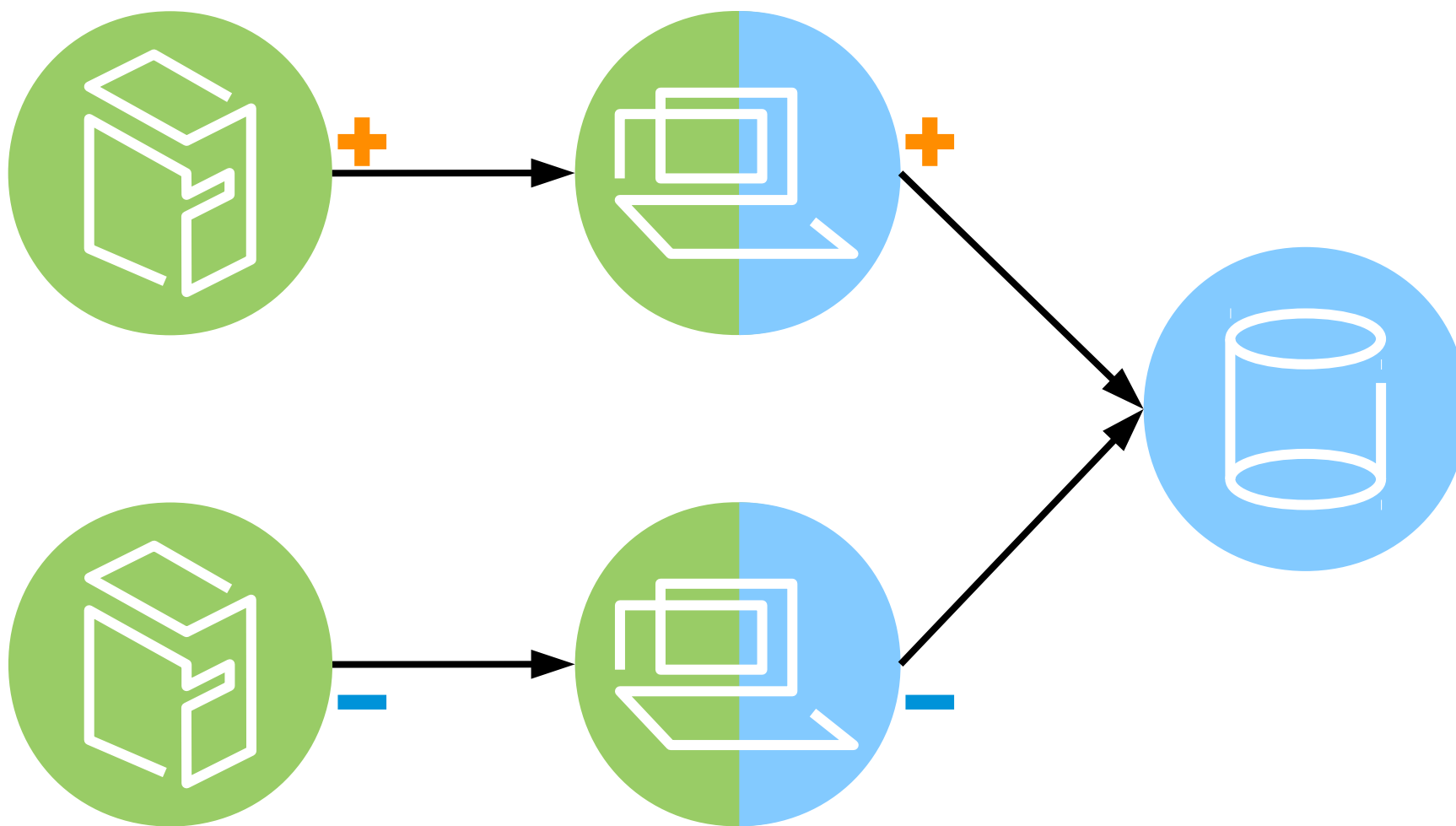
# Changelog processing



# Before









# After



Thank you